Color Constancy under Natural and Artificial Illumination

MARCEL P. LUCASSEN*†, JAN WALRAVEN†

Color constancy was studied under conditions simulating either natural or extremely artificial illumination. Four test illuminants were used: two broadband phases of daylight (correlated color temperatures 4000 and 25,000 K) and two spectrally impoverished metamers of these lights, each consisting of only two wavelengths. A computer controlled color monitor was used for reproducing the chromaticities and luminances of an array of Munsell color samples rendered under these illuminants. An asymmetric haploscopic matching paradigm was used in which the same stimulus pattern, either illuminated by one of the test illuminants, or by a standard broadband daylight (D_{65}), was alternately presented to the left and right eye. Subjects adjusted the RGB settings of the samples seen under D_{65} (match condition), to match the appearance of the color samples seen under the test illuminant. The results show the expected failure of color constancy under two-wavelengths illumination, and approximate color constancy under natural illumination. Quantitative predictions of the results were made on the basis of two different models, a computational model for recovering surface reflectance, and a model that assumes the color response to be determined by cone-specific contrast and absolute level of stimulation (Lucassen & Walraven, 1993). The latter model was found to provide somewhat more accurate predictions, under all illuminant conditions. Copyright © 1996 Elsevier Science Ltd.

Color vision   Color constancy   Color rendering   Computational vision   Cone-specific contrast   Surface reflectance

INTRODUCTION

In this paper we report experiments in which we compared the visual system's response to computer simulations of Munsell chips illuminated by either broadband light, or light composed of only two wavelengths. This was done in the context of color constancy, the ability to perceive object colors as fairly stable, independent of the spectral composition of the illuminant. In most studies of color constancy it is customary to employ a more or less "natural" illuminant–object interaction, even when the colors are simulated on a color monitor. Usually, the stimulus pattern consists, then, of Munsell chips illuminated by incandescent light or different phases of daylight (e.g. Arend & Reeves, 1986; Arend et al., 1991; Foster et al., 1992; Ho et al., 1990; Tiplitz-Blackwell & Buchsbaum, 1988).

The reason why we chose to also measure color constancy under extremely impoverished spectral conditions is twofold. First, we wanted to further test the applicability of an undoubtedly too simple, but so far

accurate, model derived in a preceding study (Lucassen & Walraven, 1993). The model in question was based on data from a rather synthetic world, characterized by a trichromatic illuminant–object interaction commonly used in computer graphics (cf. Borges, 1991). The present study provides "real world" data, obtained under conditions employing a realistic illuminant–object interaction, both for natural and artificial illuminants.

Our second reason for doing these experiments is the need for experimental tests of a fairly recent class of computational models of color constancy (e.g. Brill & West, 1986; Buchsbaum, 1980; D'Zmura & Lennie, 1986; Forsyth, 1990; Maloney, 1986, 1992; Maloney & Wandell, 1986; van Trigt, 1990). These models typically aim at recovering the spectral information that is lost in the process of light absorption in the eye's photopigments. This implies decomposing the light reflected from a surface, into its two constituent spectral distributions, i.e. the spectral power distribution of the illuminant and the reflectance function of the surface in question. The underlying principle used for such a spectral dissociation relies on the spectral constraints that have been found to hold for our natural environment. It can be shown, by principal component analysis, that the spectral power distribution of phases of daylight can be approximated by only three basis functions (Judd et al., 1964). A similar simplification can be applied to surface reflectances

*To whom all correspondence should be addressed at: Technology Center Colorimetry, Akzo Nobel Coatings b.v., P.O. Box 3, 2170 BA Sassenheim, The Netherlands.
†TNO Human Factors Research Institute, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands.

(Cohen, 1964), for which three basis functions may also account for most of the variance (Dannemiller, 1992; Maloney, 1986). Given the two sets of basis functions and an estimate of the color of the illuminant (in terms of CIE or receptor coordinates), the latter can be eliminated (e.g. Buchsbaum, 1980), and hence, surface reflectance extracted. The estimation of the illuminant is usually obtained indirectly, e.g. by taking samples of reflected light from a sufficiently large collection of surface reflectances (Buchsbaum, 1980; Maloney & Wandell, 1986). This is, in a nutshell, the rationale underlying the majority of the recent (linear) computational approaches to color constancy. For a more detailed discussion, see the comprehensive introductions by D'Zmura and Lennie (1986) or by Thompson et al. (1992).

For flat, homogeneously illuminated surfaces, and within the spectral constraints of naturally occurring surface reflectance functions and illuminant spectral power distributions, the aforementioned computational models should be quite successful in recovering surface reflectance, and hence, be capable of good color constancy. When these preconditions are not met, the models may be expected to fail, of course. However, such failures should be precisely predictable, for a given choice of model and illuminant–surface interaction (Maloney, 1992). Therefore, as a first step in the validation of this class of models, it would be informative to compare model predictions and experimental data under both favorable and adverse illuminant conditions.

Although the primary goal of this study is to show the general applicability of our earlier data analysis (that is, without having to consider spectral constraints) we shall also present predictions that are obtained by a computational model based on the principles outlined above. We shall refer to this model as the "Judd–Cohen model", since it incorporates the linear approximations of illuminant and reflectance spectra as reported by Judd et al. (1964) and Cohen (1964), respectively.

In this study, as in most other studies on color constancy, we only address the purely sensory aspect of color vision. The subjects are asked to match the color and brightness of samples seen under different illuminants. This task can be performed with good reproducibility (Lucassen & Walraven, 1993), requires no long training sessions and can be shown to yield a relatively high degree of color constancy. Other methods might have been used as well (see the Discussion), but since we wanted to test the applicability of the model derived in our previous study, we decided to stick to the same method.

METHODS

General outline of experimental method

Subjects saw two displays, which we call "test" and "match", alternately with the left and right eyes. Each display simulated an identical array of 35 Munsell chips on a neutral background. On the "test" display, seen by the left eye, four different test illuminants were used. In

TABLE 1. Munsell renotations and CIE x, y, β equivalents (under D_{65} white light) of the 30 chromatic and six achromatic samples of the stimulus pattern shown in Fig. 1

Sample number in Fig. 1	Simulated Munsell chip	x, y, β equivalents under illuminant D_{65}		
		x	y	β
1	10 YR 5/2	0.3579	0.3637	0.1939
2*	5 PB 5/4	0.2733	0.2897	0.1997
3	10 G 5/2	0.2964	0.3489	0.1924
4*	5 P 5/4	0.3023	0.2877	0.1967
5*	5 G 5/4	0.2875	0.3804	0.1969
6	10 Y 5/2	0.3460	0.3783	0.1900
7*	5 B 5/4	0.2548	0.3099	0.1946
8	10 R 5/6	0.4398	0.3604	0.1858
9	10 Y 5/6	0.4113	0.4769	0.1901
10	N 3.5/	0.3151	0.3303	0.0881
11	10 GY 5/2	0.3155	0.3684	0.1921
12	10 RP 5/6	0.3916	0.3151	0.1903
13	N 6.5/	0.3139	0.3308	0.3635
14	10 YR 5/6	0.4462	0.4244	0.1962
15	10 GY 5/6	0.3175	0.4494	0.1938
16*	5 BG 5/4	0.2633	0.3432	0.1925
17	10 R 5/2	0.3527	0.3425	0.1946
18*	N 5.0/	0.3146	0.3318	0.1983
19	10 BG 5/2	0.2844	0.3268	0.1963
20*	5 R 5/4	0.3802	0.3346	0.1931
21	10 P 5/6	0.3297	0.2782	0.1920
22	10 B 5/6	0.2310	0.2739	0.1992
23	N 2.5/	0.3147	0.3314	0.0458
24	10 G 5/6	0.2541	0.3797	0.1986
25	10 P 5/2	0.3202	0.3130	0.1909
26	N 6.0/	0.3140	0.3308	0.3068
27	10 PB 5/6	0.2733	0.2573	0.2021
28	10 BG 5/6	0.2247	0.3144	0.2000
29*	5 YR 5/4	0.4037	0.3749	0.1898
30	10 PB 5/2	0.2991	0.3044	0.1965
31*	5 RP 5/4	0.3494	0.3094	0.1895
32*	5 GY 5/4	0.3525	0.4256	0.1915
33	10 RP 5/2	0.3400	0.3272	0.1965
34*	5 Y 5/4	0.3965	0.4203	0.1917
35	10 B 5/2	0.2862	0.3131	0.1977
36	N 7.0/	0.3138	0.3312	0.4359

β represents luminance reflectance relative to an ideal white reflector (BaSO$_4$). The 11 samples of the test set are indicated by an asterisk.

the "match" display, seen by the right eye, the subject adjusted the central patch to match that of the test display. The illumination on this display was always D_{65} daylight. The displays were seen alternately, for 5 sec each, with a brief dark interval in between (the switching time for the shutter). Conditions differed as to which of the four test illuminants was used, and which Munsell test sample, chosen from a subset of 11 out of the 35, was placed in the center of the test display to be matched. This made 44 conditions in all. The subjects ran four sessions, each session dealing with one of the four test illuminants.

A more detailed account, from stimulus preparation to observer's task, is given below.

Surface reflectances

The spectral reflectance, $R(\lambda)$, of 36 samples from the Munsell Book of Color (glossy finish) were measured in
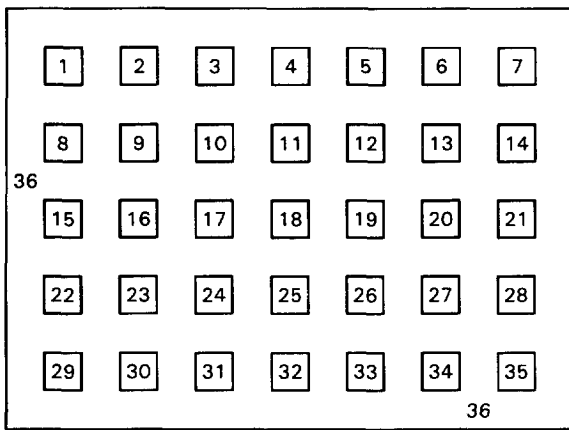
FIGURE 1. Stimulus geometry. The 1.3 deg squares are separated by a 1.3 deg grid. The background (grid) dimensions are 19.5 × 14.3 deg, somewhat smaller than the whole monitor screen (about 20 deg). The remaining area on the screen was black. See Table 1 for colorimetric specifications of the numbered samples.

the range $390 \leqslant \lambda \leqslant 730$ nm at 2-nm wavelength intervals with a SpectraScan PR-702AM spectroradiometer (Photo Research). The reflectances were measured relative to a $BaSO_4$ white, in the 0/45 deg measuring geometry. The CIE x, y chromaticities and luminance factor β (relative to white) of these samples under various illuminations, $E(\lambda)$, were computed by first calculating the X, Y, Z tristimulus values, using the numerical procedure:

$$X = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{x}(\lambda)\Delta\lambda \tag{1}$$

$$Y = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{y}(\lambda)\Delta\lambda \tag{2}$$

$$Z = \sum_{\lambda=390}^{730} E(\lambda)R(\lambda)\bar{z}(\lambda)\Delta\lambda \tag{3}$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ represent the CIE 1931 color matching functions and $\Delta\lambda = 2$ nm. The colorimetric specifications of the 36 Munsell samples under illuminant D_{65} are listed in Table 1. We used 30 chromatic and six achromatic samples, presented as a 5 × 7 matrix of square patches on a homogeneous background (one of the six achromatic samples). This was the same stimulus pattern as used in our earlier studies on color constancy and chromatic induction (Lucassen & Walraven, 1993; Walraven et al., 1991). The chromatic samples were selected from three loci of equal Munsell Chroma (/6, /4 and /2) at Munsell Value 5/, the neutrals ranged from Value 2.5 through 7.0. The samples were presented on a neutral background ($n = 7.0/$), resulting in a relative reflectance (sample to background) of 46%. The numbers of the samples in Fig. 1 correspond to those in Table 1. Eleven samples (ten chromatic and one neutral), indicated by an asterisk in Table 1, were used as test stimuli in our matching paradigm.

As discussed in our preceding paper (Lucassen &

Walraven, 1993), the distribution of the samples over the stimulus array was not random. It ensured a more or less balanced average color, locally (averaged over neighboring patches) as well as globally. The variety in color samples ensured an adequate sampling of color space. It is also in compliance with the requirements for computational models of color constancy that typically depend on an adequate number of surface reflectances in order to obtain a good estimate of the illuminant (e.g. Maloney & Wandell, 1986).

Illuminants

Two classes of illuminants were simulated: three (natural) broadband daylights and two (artificial) two-wavelength compositions. One of the broadband illuminants (D_{65}) was used for illuminating the match (reference) pattern, the other four served as test illuminants for the test pattern.

The relative spectral radiant power distributions of the three broadband illuminants were generated by the CIE method—derived from the principal components analysis of Judd et al. (1964)—as described in Wyszecki and Stiles (1982). This method takes as input the correlated color temperature (T_c) of a daylight illuminant D, where T_c may range from 4000 to 25,000 K. The output is a spectrum $E(\lambda)$, with λ in steps of 10 nm. In order to obtain the same spectral resolution as in the reflectance measurements (2 nm) we interpolated $E(\lambda)$ at 2 nm intervals.

In our simulation, the standard (white) illuminant D_{65} ($T_c = 6500$ K, $x = 0.3127$, $y = 0.3290$) was used for illuminating the match (reference) pattern. The two other daylight illuminants, D_{40} ($T_c = 4000$ K, $x = 0.3823$, $y = 0.3838$) and D_{250} ($T_c = 25,000$ K, $x = 0.2499$, $y = 0.2548$), were used as broadband test illuminants. Strictly speaking, the CIE method for generating the spectral power distribution of daylight illuminants requires the illuminant's x-coordinate to satisfy $0.25 < x < 0.38$. The x-chromaticities of D_{40} and D_{250} (0.3823 and 0.2499) violate these boundary conditions, but the violations are so small that we may safely assume that this does not affect the reality aspect of our simulation.

The other two test illuminants, designated by M_{40} and M_{250} (using M for metameric), were each composed of two wavelengths, λ_1 and λ_2. For M_{40}, $\lambda_1 = 592$ nm and $\lambda_2 = 491.8$ nm, and for M_{250}, $\lambda_1 = 560$ nm and $\lambda_2 = 433.7$ nm.* The relative intensities (power ratio $I_{\lambda_2}/I_{\lambda_1}$) of these wavelengths were 1.566 for M_{40} and 1.254 for M_{250}, so as to yield the same x, y chromaticities of illuminant M_{40} and M_{250} as for D_{40} and D_{250}. Thus, M_{40} was metameric with D_{40}, and M_{250} was metameric with D_{250}. The intensity of the (homogeneous) illuminants was such that a perfectly reflecting white diffuser would have a luminance of 30.4 cd/m², resulting

---

*In order to compute X, Y, Z tristimulus values according to Eqs (1)–(3) we interpolated the color matching functions and the reflectance spectra at 0.1 nm steps and used $\Delta\lambda = 0.1$ nm for these two illuminants.

in a luminance of the chromatic samples (Value 5) of about 6 cd/m^2 under D_{65}, consistent with our earlier studies.

The x, y chromaticities of the Munsell samples under the broad band illuminants, D_{40} and D_{250}, are those found when viewing the Munsell Book in outdoor illumination (ignoring atmospheric effects etc.), that is, they are realistic (natural) values. The x, y chromaticities of the Munsell samples rendered under the two-wavelengths lights, M_{40} and M_{250}, fall on the lines that connect the corresponding wavelengths in CIE x, y chromaticity space. Although rather unnatural, such stimuli are physically realizable in the laboratory by using laser lights, narrow-band interference filters or monochromators.

Stimulus presentation

The x, y, Y equivalents of the samples under the various illuminants were displayed on a calibrated high resolution color monitor (Sony, 1152 × 900 pixels) that was controlled by a Sun 3/260 computer (24 bit/color). For the human eye, the video RGB metamers are physically indistinguishable (as far as color is concerned) from their paper counterparts. The calibration procedure for the monitor, and the colorimetric equations required for displaying specified x, y, Y values on a color monitor, have been published elsewhere (Lucassen & Walraven, 1990).

In each experimental condition, two displays were used: a test pattern, i.e. the samples as arranged in Fig. 1 under one of the test illuminants D_{40}, D_{250}, M_{40}, or M_{250}, and a match pattern of identical geometry, illuminated by D_{65}. A pyramidal box (1 m length) with two viewing holes was placed in front of the monitor. A mechanical shutter system, located just behind the two viewing holes, alternately occluded the left and right viewing hole. In this way, each eye was locked to one or the other of the two successive illuminant conditions (test or match) to be compared. The colors of the test and match pattern were changed during the switching time of the shutters, which only took a fraction of a second. The presentation time of each pattern was set at five seconds. This was long enough for the stimulus to "settle" (at these relatively low light levels) and short enough not to disrupt the comparison of test and match sample.

Procedure

After about 5 min of dark adaptation and a few more minutes for adapting to the average luminance and color of the test pattern, the observer started the first presentation of the two alternating illuminant conditions. When viewing the test (left eye) and match pattern (right eye) the observer concentrated on the central patch. The color of the matching sample, which was initially black, was under mouse control. Each mouse movement was translated by the computer into a movement through CIE x, y color space, after which the color of the matching sample was updated accordingly. Two of the three mouse buttons were pressed to increase or decrease the

luminance of the patch at constant x, y chromaticities. The third mouse button was pressed to indicate that a satisfactory match had been obtained, after which the next test patch was presented (in total 11 samples, in pseudo-random order).

Even for unexperienced subjects, this matching procedure was easy to comprehend and required only a few training sessions to obtain reliable results. In our previous study (Lucassen & Walraven, 1993) we reported on a pilot experiment in which the test and match illuminants were identical (D_{65} white). That experiment was performed to test the reliability of the experimental method, and in particular the precision with which a haploscopic color match can be made. For the set of 11 test colors (the same set as used in the present study) the average chromatic deviation between test color and observer match was $\bar{\Delta}_{xy} = 0.008$. Compared to the size of the chromatic shifts measured in color constancy, that precision is sufficient to allow gathering data without repetition. Therefore, each subject made only one match per sample per illuminant condition. Two naive observers (AV and EG) and the first author (ML), all with normal color vision, served as subjects.

Task

The observers adjusted the central patch in the match pattern to make it match the perceived hue, saturation, and brightness of the corresponding sample in the test pattern. They were free to make eye movements and to use as many test/match alternations as were necessary to obtain a satisfactory match. All three observers reported that they were able to set satisfactory matches.

Data analysis

Chromatic constancy index. Data from experiments on color constancy may exhibit more or less constancy, depending on the experimental paradigm used. Arend *et al.* (1991) introduced a chromatic constancy index for expressing the degree of color constancy that they obtained in their experiments on, what they called, "simultaneous color constancy". This index, I, which is essentially a chromatic Brunswik ratio (Brunswik, 1928) is defined as

$$I = 1 - \frac{b}{a} \qquad (4)$$

where a and b represent Euclidean distances in CIE 1976 $u'v'$ color space. When applied to our data, a and b can be computed from

$$a = \left((u'_{\square} - u'_{o})^2 + (v'_{\square} + v'_{o})^2 \right)^{1/2} \qquad (5)$$

$$b = \left((u'_{\bullet} - u'_{o})^2 + (v'_{\bullet} + v'_{o})^2 \right)^{1/2} \qquad (6)$$

where the symbols in the subscripts refer to the symbols we used for labeling our data (see Fig. 2). That is, open squares for the samples under test illumination, open circles for the samples under match illumination, and solid circles for the matches to the test samples. In case of
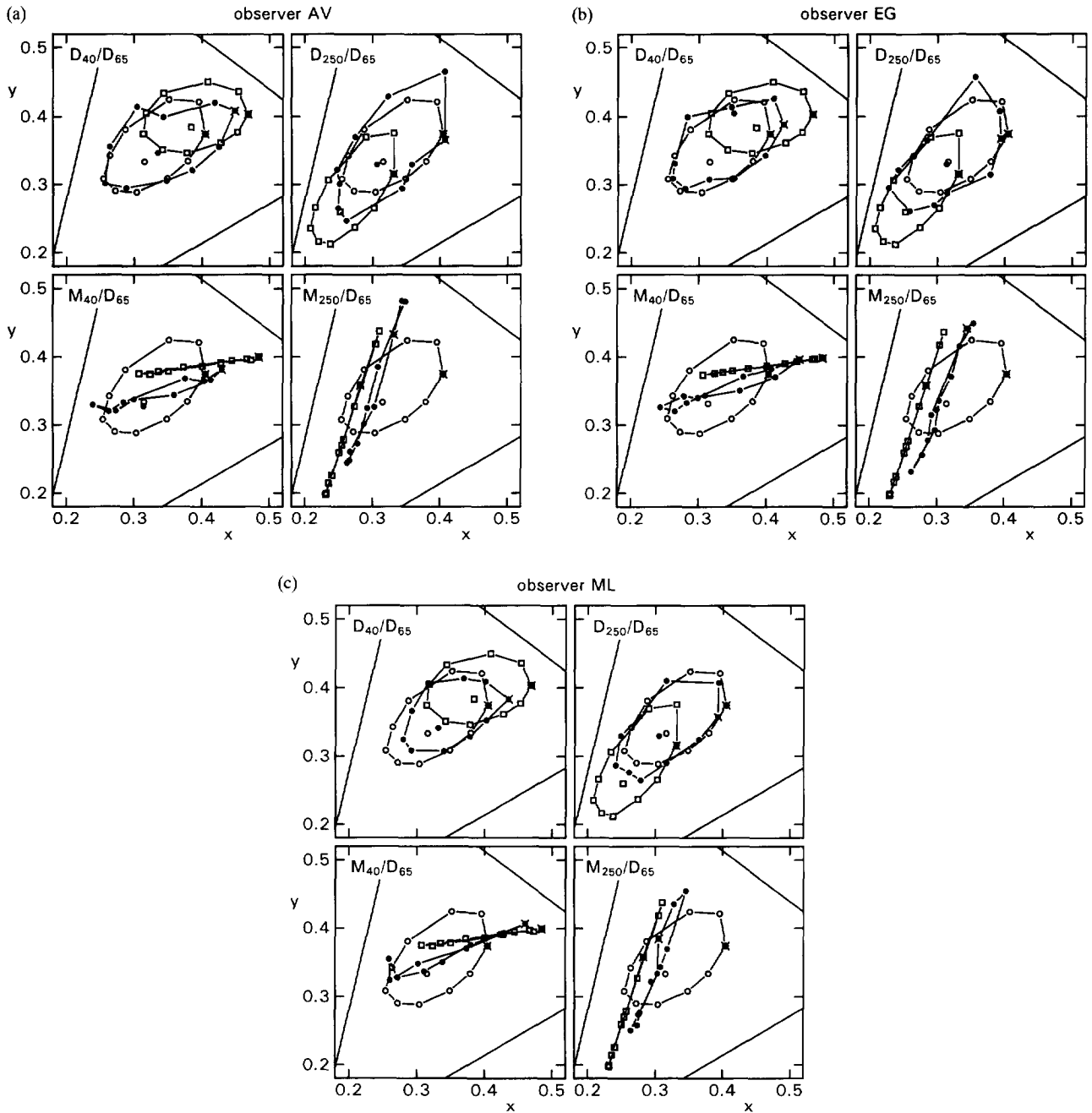
FIGURE 2. Experimental results with the daylight illuminants (top panels) and the two-wavelength illuminants (bottom panels) for obs. AV (a), EG (b) and ML (c). $\bigcirc$, Chromaticities of the test samples under D_{65}; $\square$, chromaticities of the test samples under the test illuminant in question; $\bullet$, chromaticities of the observer's matches (under D_{65}) to the test samples under the test illuminant; $\times$, test sample No. 29 (see Table 1 also). Taking this sample as the starting point, in clockwise rotation the following test samples are found along each locus: 34, 32, 5, 16, 7, 2, 4, 31, and 20.

perfect color constancy, the solid and open circles in Fig. 2 would coincide, and the same would be true for the $u'v'$-transformed versions of Fig. 2. Consequently, b would be zero, and the constancy index $I = 1$. In the opposite case (no color constancy), the solid circles would coincide with the open squares. This would imply $a = b$, and hence, $I = 0$.

Since our study deals with both experimental and predicted data, the chromatic constancy index will be applied to both sets of data. In the case of predicted data

the index is computed in exactly the same way, but with the observer matches replaced by the predicted matches. The two indices will be called I_e and I_p, for referral to experimental and predicted data, respectively.

Prediction error. A model may be theoretically capable of achieving perfect color constancy ($I_p = 1$), but that does not necessarily make it a valid model for the visual system. We therefore also computed a performance measure, the prediction error ($\bar{\Delta}_{u'v'}$), which corresponds to the mean chromatic difference between

the predicted and experimentally obtained chromaticity matches. The prediction error for the 11 test samples is computed by

$$\bar{\Delta}_{u'v'} = \frac{1}{11} \sum_{i=1}^{11} \left((u'_{pred,i} - u'_{exp,i})^2 + (v'_{pred,i} - v'_{exp,i})^2 \right)^{1/2}.$$

$$(7)$$

The prediction error is a better measure of the validity of a model, of course, than its predicted color constancy index.

RESULTS

Each subject made 11 color matches for each of the four test illuminant conditions. These we shall refer to as D_{40}/D_{65}, D_{250}/D_{65}, M_{40}/D_{65} and M_{250}/D_{65}, to indicate the test/match illuminant combination. In Fig. 2 the matches for the separate observers are plotted in the CIE x, y diagram.

The top panels in Fig. 2 relate to the two conditions with the broadband daylight illuminants, the bottom panels to the conditions with the two-wavelengths illuminants. The straight lines in the plots represent the boundaries of the triangular color space covered by the phosphors of our CRT. Open squares represent the chromaticities of the 11 test samples under the test illuminant, open circles those under the (D_{65}) match illuminant. Points representing chromatic samples are connected by lines. The remaining isolated points represent the neutral sample. When comparing top and bottom panels, note the difference in the chromaticities of the colors rendered under the test illuminant. The solid circles indicate the chromaticities of the observer's matches to the test samples.

The physical effect of changing the illuminant from broadband (upper panel) to two-wavelengths illuminants (lower panel), is to collapse the chromaticity locus of the test colors onto a straight line (see the open squares). This is the line connecting the chromaticities of the two wavelengths of the M_{40} or M_{250} light source. Under monochromatic light, all chromaticities would project onto a single point.

In Fig. 2 perfect color constancy would be indicated by coinciding solid and open circles, but this is never the case. As expected, the deviations from perfect constancy are smaller for the daylight illuminants than for the two-wavelengths illuminants. There is a general tendency for the match to the neutral test sample to be shifted back in the direction of the neutral chromaticity of that sample under white light (the solid circle in the middle). Such a shift is in accordance with an incomplete von Kries color transformation scheme (von Kries, 1905). The chromatic test samples are shifted back in the same direction, but for conditions M_{40}/D_{65} and M_{250}/D_{65} the loss of the original chromaticity spacing cannot be undone. For all observers, the color matches fall on a single line (within experimental spread). That line is translated (and rotated for condition M_{40}/D_{65}), away from the line that connects the physical chromaticities under the test illuminant.

TABLE 2. Maximum distances, in x, y space, between individual (Fig. 2) and averaged (Fig. 3) observer matches

Sample number	D_{40}/D_{65}	D_{250}/D_{65}	M_{40}/D_{65}	M_{250}/D_{65}
2	0.011	0.010	0.013	0.010
4	0.020	0.022	0.007	0.009
5	0.020	0.007	0.016	0.013
7	0.019	0.013	0.022	0.019
16	0.024	0.006	0.004	0.013
18	0.042	0.005	0.020	0.015
20	0.017	0.014	0.017	0.013
29	0.019	0.010	0.022	0.041
31	0.023	0.019	0.023	0.015
32	0.016	0.035	0.006	0.036
34	0.013	0.039	0.021	0.026
Mean	0.020	0.016	0.015	0.019

The first column contains the numbers of the 11 test samples, indicated by an asterisk in Table 1. The bottom row shows the mean maximum distance, averaged over the 11 test samples.

In the Data Predictions we show that both the moderate and gross violations from perfect color constancy are actually predicted by assuming that the visual system responds to cone-specific contrast. Before presenting these predictions, however, we shall have a closer look at the differences between observer matches and the degree of color constancy exhibited by the data shown in Fig. 2.

Table 2 presents, for each test sample and illuminant condition, the maximum individual deviations (Euclidean distance) from the averaged x, y chromaticities (of the matches). These deviations are expressed in x, y units, so a value of 0.01 in Table 2 means that the individual x, y matches for that particular sample are located within a circle of radius 0.01 centered on the average x, y value. The data in Table 2 show that, on average, the maximum distance lies between 0.015 and 0.020.

We computed the (experimental) constancy index I_e, averaged over the 11 test samples, for each observer and each illuminant condition. The mean values and standard deviations of I_e are presented in Table 3. The mean constancy index ranges from about 38% (condition M_{250}/D_{65}) to 76% (condition D_{250}/D_{65}), while standard deviations range from 0.1 to 0.2. As expected, the index for the two-wavelengths test illuminants is smaller than that for the daylight test illuminants, indicating less color constancy.

When comparing the constancy index values in Table 3 with the values that Arend et al. (1991) found for their "unasserted color matches" (what we call sensory matches), it appears that our data reveal a higher degree of color constancy (some 20% in terms of the constancy index). Troost and de Weert (1991b), who studied color constancy with both successive and simultaneous stimulus presentation, did not find large differences in a two-dimensional Brunswik ratio (a comparable constancy index) for their "exact matching" conditions. Their results were in agreement with the results of the Arend et al. (1991) study, in the sense that they too obtained a relatively low degree of color constancy.

TABLE 3. Chromatic constancy index averaged over the 11 test samples, for each observer and illuminant condition; these are the indices based on the experimental data (I_e), as computed with Eq. (4)

Observer	D_{40}/D_{65}		D_{250}/D_{65}		M_{40}/D_{65}		M_{250}/D_{65}	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
AV	0.586	0.190	0.723	0.169	0.561	0.181	0.378	0.201
EG	0.741	0.224	0.757	0.112	0.532	0.140	0.442	0.212
ML	0.600	0.103	0.750	0.113	0.469	0.134	0.417	0.206

TABLE 4. Luminance ratio of the test and match sample, averaged over the 11 test samples, for each observer and illuminant condition

Observer	D_{40}/D_{65}		D_{250}/D_{65}		M_{40}/D_{65}		M_{250}/D_{65}	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
AV	1.149	0.049	0.970	0.048	1.087	0.076	0.984	0.056
EG	1.011	0.059	1.007	0.041	1.057	0.075	0.998	0.052
ML	1.067	0.048	1.007	0.046	1.043	0.042	1.041	0.063

The difference between the degrees of color constancy as measured in our experiments and those of the other two studies mentioned, may be explained in terms of chromatic adaptation. In our experiments, the two illuminant conditions were viewed alternately with the left and right eye, so that each eye was adapted to its own illuminant color. In the studies of Arend et al. (1991) and Troost and de Weert (1991b) using binocular vision, the two eyes were adapted to both the test stimulus and the match stimulus, either with simultaneous presentation of the two illuminant conditions, or with successive presentation. In a study of Eastman and Brecher (1972) that specifically addressed the effect of viewing condition on chromatic adaptation (matching of Munsell chips illuminated by lights of 3000 and 6000 K, respectively) it was found that successive haploscopic matching, the method used in this study, yielded better color constancy than the condition with binocularly viewed test and match stimulus.

The constancy index, as defined by Eq. (4), is only informative about chromatic shifts in the $u'v'$ plane, but does not relate to the luminance component. Therefore, we also computed the ratio of the test sample luminance to the match sample luminance. Mean values and standard deviations of this ratio are presented in Table 4. The data in this table show that, on average, the luminance of the match equals the luminance of the test sample. Since the luminance of the backgrounds was the same in both eyes, this implies that the luminance contrasts of test and matching samples were also the same.

DATA PREDICTIONS

In this section we compare predictions of the experimental data on the basis of two models. The first model is that derived in our previous study (Lucassen & Walraven, 1993). We shall refer to this model as "response function". The second model uses the quite different approach of recovering surface reflectance. That model will be referred to as the "Judd–Cohen model".

Response function

The results from the preceding study, which were obtained with the same stimulus configuration (but situated in an "RGB world"), could be described by the response function

$$R^P = (Q_w^p)^r \log\left(4.35 \frac{Q_j^p}{Q_w^p}\right) \quad p = L, M, S. \quad (8)$$

where Q^P represents the quantum catch per cone class, as denoted by the superscript p. Additional subscripts j and w, indicate the input from test sample and (white) background, respectively. The spectral reflectance of the latter is flat, so the background conveys the chromaticity of the illuminant. The exponent r is observer dependent $(r \approx 0.3)$. The response function presented in Eq. (8) has to be applied to both the test and match eye. The
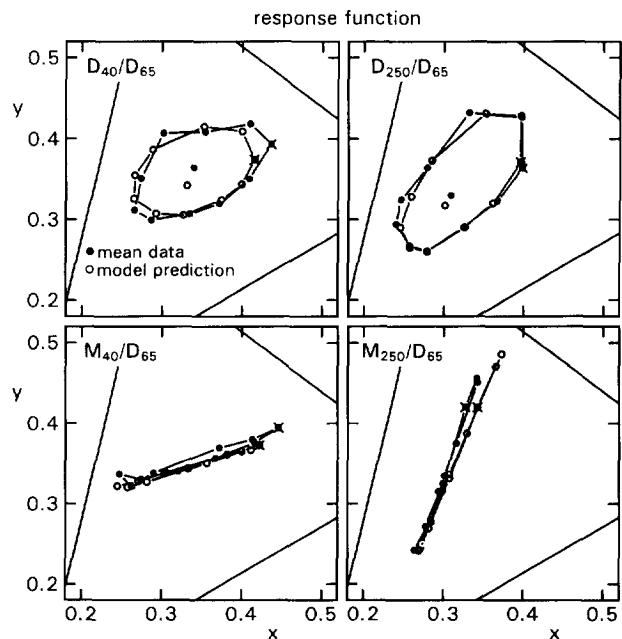

FIGURE 3. Mean observer matches (●) and predictions (○) based on the response function derived by Lucassen and Walraven (1993). Predictions computed with Eq. (9) for $r = 0.33$.

prediction of the match, in terms of cone inputs (Q_j^p), is obtained by equating the test and match eyes responses, $R^{p,t} = R^{p,m}$, where superscripts t and m denote test and match, respectively. These superscripts have to be applied to each element in Eq. (8). The cone inputs required for the matching sample, $Q_j^{p,m}$ can be computed by substitution of Eq. (8) into $R^{p,t} = R^{p,m}$. One can thus derive

$$\log(Q_j^{p,m}) = \left(\frac{Q_w^{p,t}}{Q_w^{p,m}}\right)^r \log\left(4.35\frac{Q_j^{p,t}}{Q_w^{p,t}}\right) + \log\left(\frac{Q_w^{p,m}}{4.35}\right).$$

$$(9)$$

The predictions that are obtained by applying Eq. (9) to the data, have been cast into terms of CIE x, y chromaticity coordinates [see Lucassen & Walraven (1993) for details]. In Fig. 3 these predicted chromaticities (open circles) are shown together with the experimentally obtained values averaged over the three observers (solid circles).

It is clear from Fig. 3 that Eq. (9) provides a good description of the results.

The values of $\bar{\Delta}_{u'v'}$ for conditions D_{40}/D_{65}, D_{250}/D_{65}, M_{40}/D_{65}, and M_{250}/D_{65}, computed in that order, were 0.0074, 0.0049, 0.0073, and 0.0072. Such small values, associated with about 95% explained data variance, were also obtained in our earlier study, in which we simulated an altogether different class and range of illuminants. Note that data from the daylight illuminants are about equally well predicted as those obtained for the two-wavelengths illuminants. Apparently, the complete disregard of spectral distribution in this analysis, does not affect the quality of the predictions, despite the fact that the illuminants differ substantially in their spectral composition.

Spectral distribution may be expected to be an important factor for models that try to estimate the illuminant. A representative of that class, to be discussed next, is what we called the "Judd–Cohen model".

The Judd–Cohen model

The Judd–Cohen model is essentially an implementation of the model of Buchsbaum (1980). It recovers surface reflectance, $R(\lambda)$, by removing the illuminant component, $E(\lambda)$, from the product of $E(\lambda)$ and $R(\lambda)$ that provides the input to the visual system. That input, for a given point in the visual stimulus, is given by the quantum catches L, M, and S:

$$L = \int E(\lambda)R(\lambda)L(\lambda)\mathrm{d}\lambda \qquad (10)$$

$$M = \int E(\lambda)R(\lambda)M(\lambda)\mathrm{d}\lambda \qquad (11)$$

$$S = \int E(\lambda)R(\lambda)S(\lambda)\mathrm{d}\lambda \qquad (12)$$

where $L(\lambda)$, $M(\lambda)$, and $S(\lambda)$ denote the spectral sensitivities of the long-, middle-, and short-wave sensitive cones, for wavelengths in the visual range
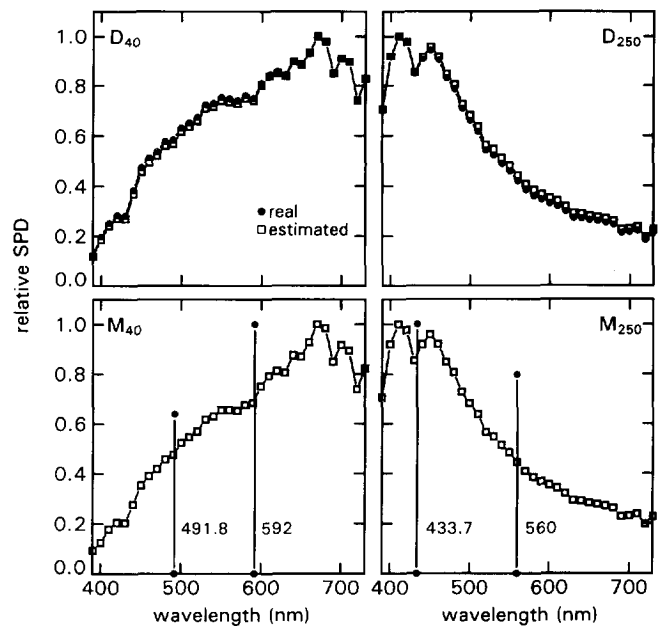

FIGURE 4. Relative spectral power distributions (SPD) of the four test illuminants ($\bullet$), compared to their estimated distributions ($\square$), as obtained with the Judd–Cohen model. D_{40} and D_{250} are broadband lights; M_{40} and M_{250} are the corresponding metamers, consisting of the two wavelengths indicated.

$390 \leqslant \lambda \leqslant 730$ nm. Using the Judd spectral basis functions, the model is capable of recovering the illuminant spectrum from only the tristimulus values (or, the linearly related quantum catches L, M, S) of the stimulus. How to obtain that trichromatic information, is the central issue in computational models of color constancy. Of the various strategies adopted, the one most frequently used was the *gray world assumption* (Evans, 1948; Buchsbaum, 1980). It implies that the illuminant can be estimated from the average chromaticity in the visual scene. Since our stimulus pattern should closely obey the gray world assumption (the samples are regularly spaced over the hue circle, and the gray background occupies about 80% of the stimulus area), we may assume that the color of the illuminant (its tristimulus values), can be accurately estimated by the model. However, the model could not "know" of course, that, in the case of the artificial illuminants, M_{40} and M_{250}, the *XYZ* values are not associated with real (broadband) daylight spectra.

The Judd–Cohen model reconstructs the spectrum of the illuminant, $E(\lambda)$, and that of the reflectance functions, $R(\lambda)$, by a linear combination of spectral basis functions:

$$E(\lambda) = e_1 E_1(\lambda) + e_2 E_2(\lambda) + e_3 E_3(\lambda) \qquad (13)$$

$$R(\lambda) = r_1 R_1(\lambda) + r_2 R_2(\lambda) + r_3 R_3(\lambda) \qquad (14)$$

in which $E_1(\lambda)$, $E_2(\lambda)$, $E_3(\lambda)$ correspond to the (first three) basis functions for the phases of daylight (Judd *et al.*, 1964), and $R_1(\lambda)$, $R_2(\lambda)$, $R_3(\lambda)$ correspond to the (first three) basis functions for reflectance spectra of Munsell chips (Cohen, 1964). The recovery of the spectral reflectance function for a given surface is simply a matter of substitution of Eqs (13) and (14) in Eqs (10)–(12) and solving the latter equations for the three basis
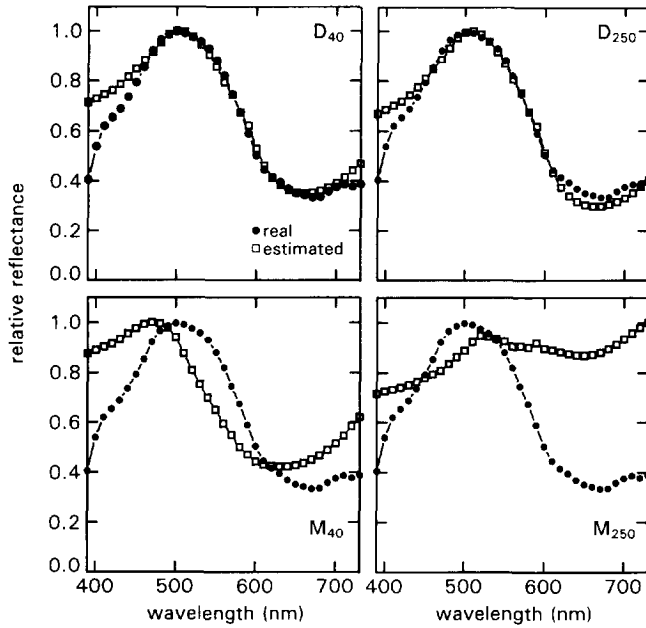
FIGURE 5. Estimated relative reflectance spectra ($\Box$), computed with the Judd–Cohen model for the same Munsell chip (5BG 5/4), but under different test illuminants (D_{40}, D_{250}, M_{40}, and M_{250}). Each panel also shows the real reflectance function in question ($\bullet$).

reconstructions of the SPDs of the broad band illuminants (top panels in Fig. 4). This is no longer the case, of course, for the two-wavelengths metamers (M_{40} and M_{250}). The x, y estimates are very similar for these lights (due to the validity of the gray world assumption in the case of our experiment), so the corresponding SPD estimates closely resemble those of the broad band lights. However, the reality is different now, the SPDs consisting only of two wavelength spikes. Consequently, the model should fail, in a predictable way, to correctly estimate reflectance. Before showing these predictions, we shall first describe how the reflectances were computed.

Recovering surface reflectance. Once the illuminant spectrum has been estimated, Eqs (10)–(12) can be solved for the three basis coefficients r_1, r_2, r_3 that determine the spectral reflectance of the surface in the stimulus pattern with inputs L, M, and S. The set of equations that have to be solved are conveniently written in vector and matrix notation:

$$\begin{pmatrix} L \\ M \\ S \end{pmatrix} = T \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \tag{15}$$

where the 3×3 matrix T is defined as

$$T = \begin{pmatrix} \int R_1(\lambda)E(\lambda)L(\lambda)d\lambda & \int R_2(\lambda)E(\lambda)L(\lambda)d\lambda & \int R_3(\lambda)E(\lambda)L(\lambda)d\lambda \\ \int R_1(\lambda)E(\lambda)M(\lambda)d\lambda & \int R_2(\lambda)E(\lambda)M(\lambda)d\lambda & \int R_3(\lambda)E(\lambda)M(\lambda)d\lambda \\ \int R_1(\lambda)E(\lambda)S(\lambda)d\lambda & \int R_2(\lambda)E(\lambda)S(\lambda)d\lambda & \int R_3(\lambda)E(\lambda)S(\lambda)d\lambda \end{pmatrix} \tag{16}$$

coefficients r_1, r_2, r_3. When r_1, r_2, and r_3 are found, the spectral reflectance function can be reconstructed with Eq. (14). However, this procedure can only succeed when $E(\lambda)$ is known; otherwise Eqs (10)–(12) contain too many unknowns and hence, too many solutions exist.

Estimating the illuminant. In the Methods we already discussed the CIE method [as described in Wyszecki and Stiles (1982) pp. 145–146] for deriving the relative spectral radiant power distributions of daylight illuminants. This involves the computation of the illuminant's basis coefficients (e_1, e_2, e_3 in our nomenclature) required for generating x, y chromaticities of the daylight illuminant. By adopting the gray world assumption, these chromaticities may be estimated from the spatially averaged x, y chromaticities in the stimulus pattern. For our test illuminants D_{40}, D_{250}, M_{40}, and M_{250} the spatially averaged (x, y) chromaticities are (0.3848,0.3867), (0.2532,0.2610), (0.4007,0.3867) and (0.2534,0.2644), respectively. Using these estimates we obtained the results shown in Fig. 4.

Each of the four panels of Fig. 4 shows an illuminant's spectral power distribution (SPD), as estimated by the Judd–Cohen model ($\Box$), together with its "real" SPD ($\bullet$). The latter were generated by the CIE method, using the actual x, y values instead of the estimated ones. The model makes use of the same method, but uses the aforementioned estimated x, y values. These are very close to the real ones, and thus yield nearly perfect

Note that all spectral functions appearing in the matrix T are known at this stage, so, the integrations can be performed. The three basis coefficients r_1, r_2, r_3 that we are looking for may then be computed from

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = T^{-1} \begin{pmatrix} L \\ M \\ S \end{pmatrix} \tag{17}$$

provided that the determinant $det(T) \neq 0$, and hence, T^{-1} (the inverse of matrix T) exists. Finally, the reflectance function $R(\lambda)$ can be reconstructed with Eq. (14).

An example of how the model succeeds in recovering surface reflectance is shown in Fig. 5.

The estimated reflectance functions shown in Fig. 5 all pertain to the same Munsell sample (5BG 5/4), but the estimates were obtained under four different illuminants. For comparison, each panel also shows the measured reflectance function. The top two panels, displaying the recovered reflectances under broad-band light, show that the Judd–Cohen model performs very well. As expected, and shown in the two bottom panels, this is no longer possible for the condition in which the sample is illuminated by the two-wavelength light of the illuminants M_{40} or M_{250}. Under M_{40} the reconstructed sample reflectance has been shifted towards blue, whereas under M_{250} the reflectance more resembles that of an achromatic sample.
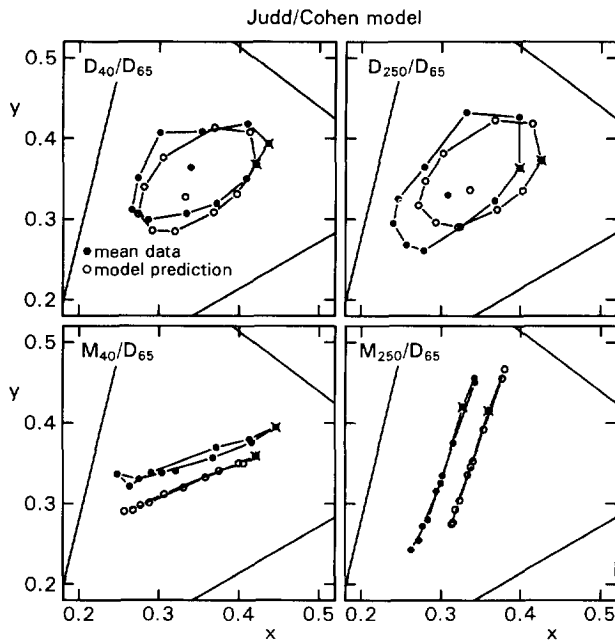
Judd/Cohen model

FIGURE 6. Mean observer matches (●) and predictions (○) of the Judd–Cohen computational model.

The results shown in Fig. 5 are of interest in that they show that a Judd–Cohen type computational model fails in a *predictable* way when applied to illuminance spectra that cannot be reconstructed with the Judd spectral basis functions. The next step is to test whether the visual system makes the same "mistakes" as those predicted by the Judd–Cohen model.

Predictions. The predictions obtained with the Judd–Cohen model, again in terms of the *x, y* chromaticities of the matching sample under the D_{65} reference illuminant, are computed by the predicted tristimulus values

$$X_{\text{pred}} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{x}(\lambda)\Delta\lambda \qquad (18)$$

$$Y_{\text{pred}} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{y}(\lambda)\Delta\lambda \qquad (19)$$

$$Z_{\text{pred}} = \sum_{\lambda=390}^{730} D_{65}(\lambda)R(\lambda)\bar{z}(\lambda)\Delta\lambda \qquad (20)$$

where $D_{65}(\lambda)$ represents the spectral power distribution of the D_{65} daylight illuminant and $R(\lambda)$ represents the recovered reflectance function for the test illuminant and sample in question.

Using the same data format as before, the predictions of the Judd–Cohen model are shown in Fig. 6.

The data predictions obtained with the Judd–Cohen model are in fair agreement with the data, but Fig. 3 shows that the performance of the response function is still better. The values for the chromatic difference (the prediction error), as defined in Eq. (7), are 0.0119, 0.0227, 0.0201, and 0.0263, for D_{40}/D_{65}, D_{250}/D_{65}, M_{40}/D_{65}, and M_{250}/D_{65}, respectively. These values are 1.6–4.6 times larger than those obtained when using the response function.

In addition to the prediction error we also computed the predicted constancy indices (I_p). This constancy index is not a criterion for evaluating the validity of the models. All it does is provide a measure for the extent to which a model is *theoretically* capable of achieving color constancy, irrespective of whether the model is correct or not. Table 5 shows the indices that the two models predict for each of the four illuminant conditions.

The constancy indices obtained for the response function are generally somewhat larger than the corresponding values obtained for the Judd–Cohen model. We did not expect this result, particularly not for the conditions with broadband illumination. The spectral distributions of D_{40} and D_{250} are quite accurately estimated by the model (see Fig. 4), so it is apparently the reconstruction of the surface reflectances that could be improved upon, possibly by using more basis functions (cf. Vrhel *et al.*, 1994). On the other hand, attempts to improve the constancy index might lead to a less realistic model of the visual system. The experimental evidence obtained so far, suggests that human color constancy is not perfect.

DISCUSSION

The experimental method used in this study, asymmetric dichoptic color matching, has frequently been used in studies on chromatic adaptation [see Wyszecki and Stiles (1982) for a review]. In as far as these studies employed stimuli representing object–illuminant interactions (as was actually not the most common practice), these studies can be classified as studies on color constancy (e.g. Burnham *et al.*, 1952). However, the associated theoretical developments that have emerged over time, bear only little resemblance to the response function we derived. Apart from methodological differences (nearly all of these studies used simultaneous instead of successive dichoptic matching) this is probably due to the different way of analyzing the data. The models often have a strong empirical character, typically

TABLE 5. Comparison of predicted chromatic constancy indices (I_p), as computed for the response function (LW) and the Judd–Cohen model (JC)

Model	D_{40}/D_{65} Mean	SD	D_{250}/D_{65} Mean	SD	M_{40}/D_{65} Mean	SD	M_{250}/D_{65} Mean	SD
LW	0.767	0.077	0.814	0.089	0.540	0.177	0.402	0.177
JC	0.651	0.070	0.757	0.051	0.474	0.352	0.402	0.398

expressed in the dimensions of CIE chromaticity space, like for example, the seven or more equations comprising the Helson–Judd formulation (Judd, 1940). Therefore, despite some affinity between our study and those earlier adaptation/constancy studies, we shall discuss our results only in the context of the more modern, computational developments in this field.

It is common practice, in both the older and more recent studies of color constancy, to employ illuminant conditions that are best suited for demonstrating the efficacy of the effect. The present study deviates in this respect by also including spectrally impoverished illuminants, the two-wavelengths metamers of D_{40} and D_{250} (M_{40} and M_{250}). By doing so we were able to measure the deterioration of color constancy, which is specifically due to the lack of spectral "capacity" of the illuminant. The particular way in which color constancy breaks down under these conditions, is informative as to how spectral information is processed by the visual system. An important issue in this respect is whether the visual system, using the Judd–Cohen "spectral tool kit", may achieve complete recovery of spectral distribution functions, as is the aim of current computational models of color constancy.

This study shows the performance of such a computational model—what we called the Judd–Cohen model—both under natural and unnatural illuminant conditions. It is not self-evident that this type of model would generate the kind of predictions we found (Fig. 6). In a qualitative sense, of course, one may expect good color constancy in the conditions with natural illumination (e.g. Brainard & Wandell, 1991), and virtually no color constancy when the light is reduced to only two wavelengths. When comparing the chromaticities of the model predictions (Fig. 6, open circles) with the chromaticities of the stimuli under D_{65} (Fig. 2, open circles), this is indeed what the predictions show. However, the central question is how accurate these predictions are borne out in the experimental data. As shown in Fig. 6, the predictions of the Judd–Cohen model are in the right direction, but there is room for improvement. It is not easy to see how this might be achieved. It is not just a matter of finding ways of boosting the constancy index. In searching for a model that would predict better color constancy, we also tested the algorithm of van Trigt (1990). As described elsewhere (Lucassen, 1993), application of this algorithm to our test stimuli yielded good color constancy for our conditions with broad-band illuminants (the latter were reconstructed with the Judd basis spectral functions). The associated values of the predicted chromatic constancy index (I_p) were 0.965 and 0.923 for D_{40}/D_{65} and D_{250}/D_{65}, which is actually better than the corresponding values obtained for the models discussed here (see Table 5). However, the good color constancy obtained with this algorithm was actually the reason for its relatively less accurate predictions of what human observers see (Lucassen, 1993). Other computational models may face the same problem, if designed to predict near perfect color constancy, as may be envisaged when using larger

numbers of (surface) spectral basis functions (Parkkinen et al., 1989; Vrhel et al., 1994).

It is conceivable that the visual system might be capable of better color constancy when measured under more natural conditions and/or with better methods. Even our "normal" (broadband) stimulus conditions might be considered as being somewhat synthetic, in the sense that the visual scene lacks a third dimension (no shadows and shading), and that the appearance of the samples is consistent with perfectly diffusing surfaces under a spatially uniform illumination. However, these are exactly some of the most important constraints—see Forsyth (1990) for a complete list—that have to be met when applying the present generation of Judd–Cohen type models of color constancy.

As for methodology, there are indeed different ways for measuring color constancy. One could argue that testing the purely sensory aspect of color perception (hue, saturation, lightness) only probes part of the underlying mechanisms. A possible alternative is to test for the correct recognition (rather than perception) of surface samples, thereby ignoring possible deviations from sensory invariance. This method was introduced by Arend and Reeves (1986), who asked their subjects to adjust the color of a match sample as if "cut from the same paper" as the test sample. The subject is thereby instructed to take into consideration that the samples are shown under different illuminants, and thus may not necessarily appear as having the same color. Subjects are apparently able to follow that instruction, thereby possibly using contextual cues, even when these are simplified to a simple disk-annulus stimulus configuration (Arend et al., 1991). However, in spite of relaxing the definition of color constancy (sensory invariance is no longer required) the latter studies did not achieve more than moderate color constancy, about 60% in terms of the chromatic constancy index (Arend et al., 1991). Other studies also show that incomplete color constancy is the rule rather than the exception (Reeves et al., 1989; Tiplitz-Blackwell & Buchsbaum, 1988; Valberg & Lange-Malecki, 1990). Considering that the visual system apparently does not strive for perfect color constancy, it would make sense to search for a mechanism that actually is intrinsically incapable of perfect color constancy. The von Kries white normalization (von Kries, 1905) has this property, and, for that reason, has sometimes been treated as inadequate for models of color constancy (e.g. Worthey & Brill, 1986; Dannemiller, 1993). The response function, which shares the trichromatic gain adjustment implied in the von Kries model, also shares the associated imperfection with respect to color constancy.

We showed that the Judd–Cohen model also predicts less than perfect color constancy. However, the way in which the predictions deviate from constancy is rather different from what the data show. There is not enough support, particularly when considering the availability of a simpler alternative, for the hypothesis that this model is implemented in the visual system. This would square

with the lack of neurophysiological evidence for structures performing the Judd–Cohen estimates of surface reflectance (D'Zmura & Lennie, 1986; Troost & de Weert, 1991a). Still, we feel that more experiments are required to test the physiological validity of computational models based on the Judd–Cohen spectral analysis. All that can be said on the basis of the present (and first) test is that there is enough reason to warn against simply accepting such models without experimental validation or considering alternative approaches.

Such an alternative is a trichromatic extension of contrast or lightness models (e.g. Hurlbert, 1986), for which the foundation was laid in the retinex model (Land, 1986; McCann et al., 1976). Our response function belongs to that class, but with some important modifications, as discussed in Lucassen and Walraven (1993). Its essential feature, responding to contrast—a sound strategy in a world where luminance varies over more than ten decades—is consistent with the results from other psychophysical studies on invariant (a)chromatic vision (e.g. Arend & Goldstein, 1990; Shapley, 1986; Wallach, 1948; Walraven et al., 1991). A system that responds to contrast can be easily implemented by a resetting mechanism or automatic gain control (Koenderink et al., 1971; Rushton, 1965; Walraven & Valeton, 1984). Such a mechanism also has the effect of removing a steady-state signal, for which there is also psychophysical evidence (Tiplitz-Blackwell & Buchsbaum, 1988; Walraven, 1976; Whittle & Challands, 1969).

As discussed elsewhere (Shapley et al., 1990; Walraven et al., 1990), the importance of contrast can also be demonstrated at the physiological level (e.g. Enroth-Cugell & Robson, 1966; Reid & Shapley, 1988; Shapley & Enroth-Cugell, 1984). As for our model's assumed cone-specificity of the contrast response, one may expect this to be reflected in receptive fields driven by single cone classes. Physiological evidence for this notion is available, but does not always exclude other interpretations. However, recently Reid and Shapley (1992) have provided unambiguous evidence for cone-specific inputs in both center and surround of parvocellular neurons.

In the section on Data Predictions, we compared predictions of the experimental data on the basis of the "Judd–Cohen model" and the "response function". The latter model, as described in Eq. (8) applies to colored reflective samples (Munsell chips) on a white background, i.e. luminance *decrements*. Although this is a condition frequently used in laboratory studies, the applicability of Eq. (8) for this condition does not imply its universal validity. As a matter of fact, in a separate study, focusing on the luminance variable in color constancy (Lucassen, 1993), Eq. (8) was shown to fail when applied to predicting data relating to luminance *increments*. In that same study a more general model was derived that incorporates luminance contrast as a separate variable, thus allowing it to describe the results obtained from stimuli in the decrement as well as the increment mode. However, Eq. (8) describes the present data as accurately as the more general (but also more complex)

model, since the latter effectively reduces to Eq. (8) for contrasts in the range $0 < C < 1$. Because of that, and also because luminance contrast was not the central issue of this study, Eq. (8) was our choice for describing the data, rather than the more elaborate formula that would be required for data varying in luminance contrast.

In conclusion, we have compared the data predictions resulting from two approaches to explaining color constancy, which differ in the way spectral information is used. We have shown that, for the limited conditions of our laboratory experiment, both the existence and breakdown of color constancy are better described by a mechanism responding to cone-specific contrast than by a system that estimates illuminant and reflectance spectra.

REFERENCES

Arend, L. E. & Goldstein, R. (1990). Lightness and brightness over spatial illumination gradients. *Journal of the Optical Society of America A, 7,* 1929–1936.

Arend, L. E. & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A, 3,* 1743–1751.

Arend, L. E., Reeves, A., Schirillo, J. & Goldstein, R. (1991). Simultaneous color constancy: papers with diverse Munsell values. *Journal of the Optical Society of America A, 8,* 661–672.

Borges, C. F. (1991). Trichromatic approximation method for surface illumination. *Journal of the Optical Society of America A, 8,* 1319–1323.

Brainard, D. H. & Wandell, B. A. (1991). A bilinear model of the illuminant's effect of color appearance. In Movshon, J. A. & Landy, M. S. (Eds), *Computational models of visual processing.* Cambridge, MA: MIT Press.

Brill, M. H. & West, G. (1986). Chromatic adaptation and color constancy: A possible dichotomy. *Color Research and Application, 11,* 196–204.

Brunswik, E. (1928). Zur Entwicklung der Albedowahrnehmung. *Zeitschrift für Psychologie, 64,* 216–227.

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute, 310,* 1–26.

Burnham, R. W., Evans, R. M. & Newhall, S. M. (1952). Influence on color perception of adaptation to illumination. *Journal of the Optical Society of America, 42,* 597–605.

Cohen, J. (1964). Dependency of the spectral reflectance curves of the Munsell color chips. *Psychonomic Science, 1,* 369–370.

Dannemiller, J. L. (1992). Spectral reflectance of natural objects: How many basis functions are necessary. *Journal of the Optical Society of America A, 9,* 507–515.

Dannemiller, J. L. (1993). Rank orderings of photoreceptor photon catches from natural objects are nearly illuminant-invariant. *Vision Research, 33,* 131–140.

D'Zmura, M. & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A, 3,* 1662–1672.

Eastman, A. A. & Brecher, S. A. (1972). The subjective measurements of color shifts with and without chromatic adaptation. *Journal of Illuminating Engineering Society, 2,* 239–246.

Enroth-Cugell, C. & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology, 187,* 517–522.

Evans, R. M. (1948). *An introduction to color.* New York: John Wiley and Sons.

Forsyth, D. A. (1990). A novel algorithm for color constancy. *International Journal of Computer Vision, 5,* 5–36.

Foster, D. H., Craven, B. J. & Sale, R. H. (1992). Immediate colour constancy. *Ophthalmic and Physiological Optics, 12,* 157–160.

Ho, J., Funt, B. V. & Drew, M. S. (1990). Separating a color signal into illumination and surface reflectance components: Theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12,* 966–977.

Hurlbert, A. (1986). Formal connections between lightness algorithms. *Journal of the Optical Society of America A, 3*, 1684–1693.

Judd, D. B. (1940). Hue, saturation and lightness of surface colors with chromatic illumination. *Journal of the Optical Society of America, 30*, 2–32.

Judd, D. B., MacAdam, D. L. & Wyszecki, G. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America, 54*, 1031–1040.

Koenderink, J. J., van de Grind, W. A. & Bouman, M. A. (1971). Foveal information processing at photopic luminances. *Kybernetik, 8*, 128–144.

von Kries, J. (1995). Die Gesichtsempfindungen. In: Nagel, W. (Ed.) *Handbuch der Physiologie des Menschen* (Vol. 3, pp. 109–282). Vieweg, Braunschweig.

Land, E. H. (1986). Recent advances in retinex theory. *Vision Research, 26*, 7–21.

Lucassen, M. P. (1993). Quantitative studies of color constancy. PhD. thesis, University of Utrecht.

Lucassen, M. P. & Walraven, J. (1990). Evaluation of a simple method for color monitor recalibration. *Color Research and Application, 15*, 321–326.

Lucassen, M. P. & Walraven, J. (1993). Quantifying color constancy: Evidence for nonlinear processing of cone-specific contrast. *Vision Research, 33*, 739–757.

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small number of parameters. *Journal of the Optical Society of America A, 3*, 1673–1683.

Maloney, L. T. (1992). A mathematical framework for biological vision. *Behavioral and Brain Sciences, 15*, 45.

Maloney, L. T. & Wandell, B. A. (1986). Color constancy: A method for recovering surface spectral reflectance. *Journal of the Optical Society of America A, 3*, 29–33.

McCann, J. J., McKee, S. P. & Taylor, T. H. (1976). Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the "color Mondrian" experiments. *Vision Research, 16*, 445–458.

Parkkinen, J. P. S., Hallikainen, J. & Jaaskelainen, T. (1989). Characteristic spectra of Munsell colors. *Journal of the Optical Society of America A, 6*, 318–322.

Reeves, A., Arend, L. E. & Schirillo, J. (1989). Color constancy in isolated displays. *Perception, 18*, 529–530.

Reid, R. C. & Shapley, R. W. (1988). Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Research, 28*, 115–132.

Reid, R. C. & Shapley, R. W. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature, 356*, 716–717.

Rushton, W. A. H. (1965). Visual adaptation. *Proceedings of the Royal Society (B), 162*, 20–46.

Shapley, R. (1986). The importance of contrast for the activity of single neurons, the VEP and perception. *Vision Research, 26*, 45–61.

Shapley, R., Caelli, T., Grossberg, S., Morgan, M. & Rentschler, I. (1990). Computational theories of visual perception. In: Spillmann, L. & Werner, J. S. (Eds), *Visual perception: The neurophysiological foundations* (pp. 53–101). San Diego: Academic Press.

Shapley, R. & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain control. In: Osborne, N. & Chader, G. (Eds), *Progress in retinal research* (p. 3). Oxford: Pergamon Press.

Thompson, E., Palacios, A. & Varela, F. J. (1992). Ways of coloring: Comparative color vision as a case study for cognitive science. *Behavioral and Brain Sciences, 15*, 1–26.

Tiplitz-Blackwell, K. & Buchsbaum, G. (1988). Quantitative studies of color constancy. *Journal of the Optical Society of America A, 5*, 1772–1780.

van Trigt, C. (1990). Smoothest reflectance functions. I. Definition and main results. *Journal of the Optical Society of America A, 7*, 1891–1904.

Troost, J. M. & de Weert, C. M. M. (1991a). Surface reflectance and human color constancy: Comment on Dannemiller (1989). *Psychological Review, 98*, 143–145.

Troost, J. M. & de Weert, C. M. M. (1991b). Naming vs matching in color constancy. *Perception & Psychophysics, 50*, 591–602.

Valberg, A. & Lange-Malecki, B. (1990). Colour constancy in Mondrian patterns: A partial cancellation of physical chromaticity shifts by simultaneous contrast. *Vision Research, 30*, 371–380.

Vrhel, M. J., Gershon, R. & Iwan, L. S. (1994). Measurements and analysis of object reflectance spectra. *Color Research and Application, 19*, 4–9.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology, 38*, 310–324.

Walraven, J. (1976). Discounting the background—the missing link in the explanation of chromatic induction. *Vision Research, 16*, 289–295.

Walraven, J., Benzschawel, T., Rogowitz, B. E. & Lucassen, M. P. (1991). Testing the contrast explanation of color constancy. In: Valberg, A. & Lee, B. B. (Eds), *From pigments to perception* (pp. 369–378). New York: Plenum Press.

Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A. & Schnapf, J. L. (1990). The control of visual sensitivity: receptoral and postreceptoral processes. In: Spillmann, L. & Werner, J. S. (Eds), *Visual perception: The neurophysiological foundations* (pp. 53–101). San Diego: Academic Press.

Walraven, J. & Valeton, J. M. (1984). Visual adaptation and response saturation. In: van Doorn, A. J., van de Grind, W. A. & Koenderink, J. J. (Eds), *Limits in perception*. Utrecht: VNU Science Press.

Whittle, P. & Challands, P. D. C. (1969). The effect of background luminance on the brightness of flashes. *Vision Research, 9*, 1095–1110.

Worthey, J. A. & Brill, M. H. (1986). Heuristic analysis of von Kries color constancy. *Journal of the Optical Society of America A, 3*, 1708–1712.

Wyszecki, G. & Stiles, W. S. (1982). *Color science, concepts and methods, quantitative data and formulae*, 2nd edition. New York: John Wiley.